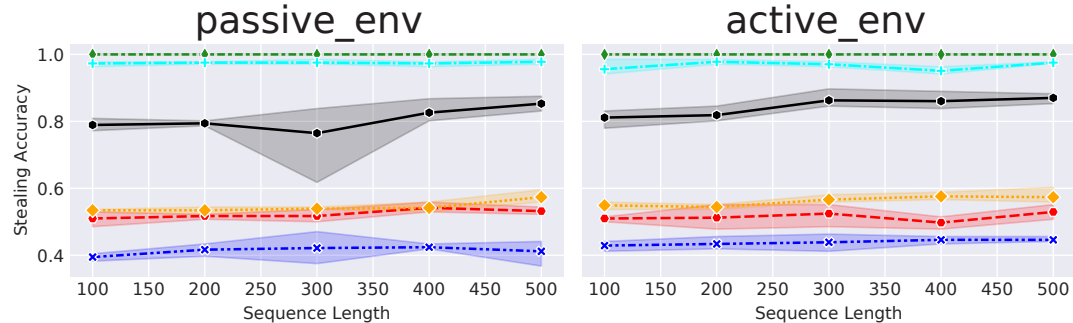
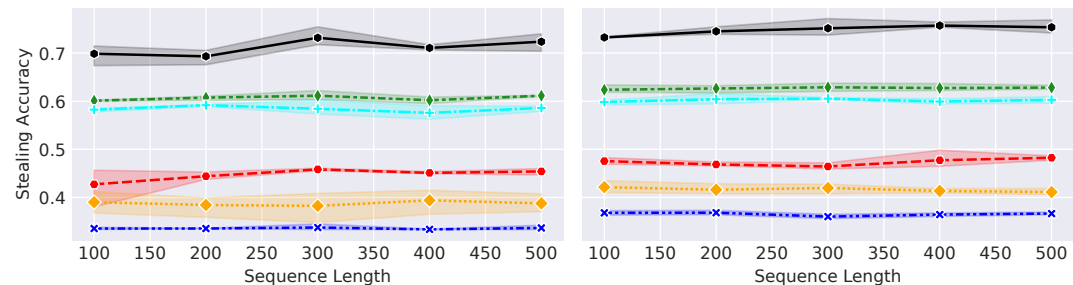


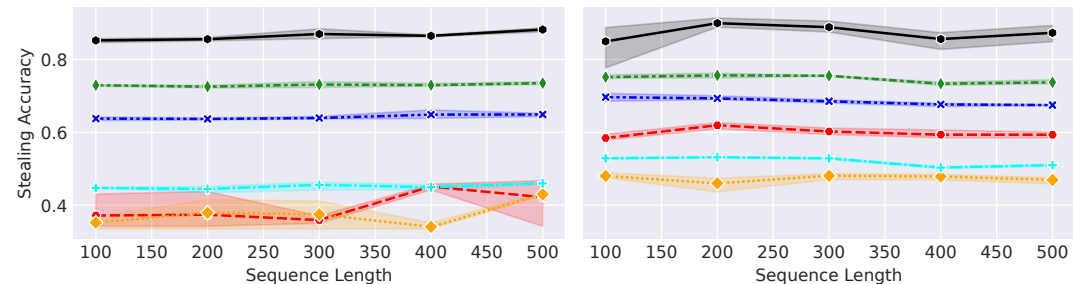
DQN



PG



PPO



activation
 hidden_layers
 neuron_number
 optimizer
 learning_rate
 gamma

Acrobot, Reward = -100

CartPole, Reward = 500